

Serial Number: 10/003,608 Changed a file from non-ASCII to ASCII

CRF Processing Date:

Edited by:

Verified by:

01/28

12/17/2001

Changed the margins in cases where the sequence text was "wrapped" down to the next line.

 ENTERED
H2

Edited a formal error in the Current Application Data section, specifically:

Edited the Current Application Data section with the actual current number. The number inputted by the applicant was the prior application data; or other _____

Added the mandatory heading and subheadings for "Current Application Data".

Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.

Changed the spelling of a mandatory field (the headings or subheadings), specifically:

Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were:

Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.

Inserted colons after headings/subheadings. Headings edited included: _____

Deleted extra, invalid, headings used by an applicant, specifically:

Deleted: non-ASCII "garbage" at the beginning/end of files; secretary initials/filename at end of file;
 page numbers throughout text; other invalid text, such as _____

Inserted mandatory headings, specifically: _____

Corrected an obvious error in the response, specifically:

Edited identifiers where upper case is used but lower case is required, or vice versa.

Corrected an error in the Number of Sequences field, specifically:

A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.

Deleted ending stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: _____

Other: Deleted duplicate 12217-12237 responses

#2 OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/003,608

DATE: 12/17/2001
TIME: 20:15:46

Input Set : A:\PTO.AMC.TXT
Output Set: N:\CRF3\12172001\J003608.raw

P5

4 <110> APPLICANT: Bartha, Gabor
5 Walker, Michael
7 <120> TITLE OF INVENTION: METHODS FOR ANALYZING GENE EXPRESSION PATTERNS
10 <130> FILE REFERENCE: ICYTP012
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/003,608
C--> 12 <141> CURRENT FILING DATE: 2001-11-01
12 <150> PRIOR APPLICATION NUMBER: 60/245,081
13 <151> PRIOR FILING DATE: 2000-11-01
15 <160> NUMBER OF SEQ ID NOS: 30
17 <170> SOFTWARE: FastSEQ for Windows Version 4.0
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 4588
21 <212> TYPE: DNA
22 <213> ORGANISM: Human
24 <400> SEQUENCE: 1
25 gccccggcca agagagatca caccggcagc cgaccctgcc agcgagcgag cccgacccca 60
26 ggcgtccatg gagcgtcgcc tccggccggc ccctggccgg accggccgct gggcgccggc 120
27 tcctgccttg accaggactt gggactttgc gaaaggatcg cggggccggg agaggtaacc 180
28 gccgcgcctc cggagaggt gttggagagc acaatggctg aacaagtctt tcctcaggct 240
29 ttgtatgttgc gcaatatgcg gaaagctgtg aagatacggg agagaactcc agaagacatt 300
30 tttaaaccta ctaatggat cattcatcat tttaaacca tgcaccgata cacactggaa 360
31 atgttcagaaa cttgcagtt ttgtcctcag tttcgggaga tcatccacaa agccctcatc 420
32 gacagaaaaca tccaggccac cctggaaagc cagaagaaac tcaactgggt tcgagaagtc 480
33 cggaaagctt gggcgttgc aacgaacggt gacggcaatt gcctcatgca tgccacttct 540
34 cagtagatgt gggcggttca gacacacgac ttggtaactga gaaaggcgct gttcagcacg 600
35 ctcaggaaa cagacacacg caactttaaa ttccgctggc aactggatc tctcaaattct 660
36 cagaaattttt ttgaaacggg gctttgtat gatactcgga actggatga tgaatggac 720
37 aatcttatca aaatggcttc cacagacaca cccatggccc gaagtggact tcagtagaac 780
38 tcactggaag aaatacacat atttgcctt tgcaacatcc tcagaaggcc aatcattgtc 840
39 atttcagaca aaatgctaag aagtttggaa tcaggttcca atttcgcctt tttgaaagtg 900
40 ggtggaaattt acttgcctct ccaactggctt gcccaggaaat gctacagata cccattgtt 960
41 ctcggctatg acagccatca ttttgcctt ttggtaaccc tgaaggacag tgggcctgaa 1020
42 atcccgagctg ttccacttgt taacagagac cggggaaagat tgaagacatt aaaagttcac 1080
43 ttttgcacag atccctgaaaa tgagatgaag gagaagctct taaaagatgat cttaatggtg 1140
44 atagaaatcc ccttccaaagg ctgggaccat ggcacaactc atctcatcaa tgccgcaaag 1200
45 ttggatgaaag ctaacttacc aaaagaaatc aatctggtag atgattactt tgaacttgg 1260
46 cagcatgagt acaagaaaatg gcaggaaaac agcgacgagg ggaggagaga ggggcacgccc 1320
47 cagaatccca tggaaaccttc cgtggcccgat ctttctctca tggatgtaaa atgtgaaacg 1380
48 cccaaactgcc ccttcttcat gtctgtgaac acccagccctt tatgcatgat gtcagatgg 1440
49 aggccggcaaa agaatcaaaa caaactccca aagctgaact ccaagccggg ccctgggg 1500
50 ctccctggca tggcgctcg ggcctctcg ggagaaggctt atgagccctt ggcgttggaa 1560
51 cctgaggagt ccactgggg gctcattcg gcccaccga cagcaccggc ccctttctg 1620
52 ttcagtgaga ccactgcccattt gaaatgcagg agccggctt gccccttac actgaatgt 1680
53 cagcacaacg gatgttgc acgttgcac aacggccggc aacttcacgc cagccacgccc 1740
54 ccagaccaca caaggcactt gatgtccggg aagtgcacag cctgcctcca gatgttacc 1800
55 aggacattt atggatctg cagtagttgc ttcaaaaagga ctacagcaga ggccttcc 1860
56 agcctcagca ccagcctccc tccttcgtt caccagcgat ccaagtcaga tccctcgccg 1920

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/003,608

DATE: 12/17/2001
TIME: 20:15:46

Input Set : A:\PTO.AMC.TXT
Output Set: N:\CRF3\12172001\J003608.raw

57 ctcgtccgga gcccctcccc gcattcttgc cacagagctg gaaacgacgc ccctgctggc 1980
 58 tgccgtctc aagctgcacg gactcctggg gacaggacgg ggacgagcaa gtgcagaaaa 2040
 59 gccggctcgcg tgtattttgg gactccagaa aacaaggggct tttgcacact gtgttcatc 2100
 60 gagtacagag aaaacaaaca ttttgcgtgc gcctcaggga aagtcaactc cacagcgtcc 2160
 61 aggttccaga acaccattcc gtgcctgggg aggaaatgcg gcacccttgg aagcaccatg 2220
 62 ttgttgcaggat actgccagaa gtgttcatc gaagctcaga atcagagatt tcatgaggcc 2280
 63 aaaaggacag aagagcaact gagatcgacg cagcgcagag atgtgcctcg aaccacacaa 2340
 64 agcacctcaa ggcccaagtg cgcggggcc tcctgcaga aacatcctggc ctgcccgc 2400
 65 gaggagctct gcatggagtgc tcagcatccc aaccagagga tggccctgg gcccacccgg 2460
 66 ggtgagctg ccccccaga ccccccagaag cagcgttgc gggcccccgc ctgtgatcat 2520
 67 ttggcaatgc ccaagtgc aacatgc aacgaatgtc ttcaatgc aacatgttat 2580
 68 ggctaaaccgg aaacagggtgg gtcacccctt gcaagaatgt gggcctcgag ctgtcaatc 2640
 69 tcatgggtct atcctctgaa cccctcagct gcccactc aactgggtt aagggtgtct 2700
 70 gacaggaga ggaaagataa gctcttcgt gtgcacatc tgctcaggat tggtaacccgg 2760
 71 ggagtgttcc caggtggctc tagaaagcaa agttgttac tggcaaggaa tgatgtcaga 2820
 72 ttccatggccaa ggttccctt ctccatccaa gcaggaggcc aggaacttct ttggacttgg 2880
 73 aagggtgtcg gggactggcc gaggccctgt cccctgcgc atcaggactg ctcatcg 2940
 74 ttggctgaga aaggaaaaag acacacaatc cgcgtgggtt ggagaagcca gagcattcc 3000
 75 acctccctcc ccccaagcatc ttcagatgc gtgaagccag atcctcatgg cagcggggcc 3060
 76 ctctgcaga aactcaagga aactcaaggaa aatggacgt attcagagat tgttttagt 3120
 77 tcatgggtt tccatccatc cccgggttcc ttccatccatc cccggcagaa atgcagaacc 3180
 78 atccatggac tttgttgcaggatgc gacttgcacat gttcacattt acagaaaaac 3240
 79 aagctgtct ttataatatc caccatccaa aaaatccatc tttttactt ggaagacgtg 3300
 80 taactctttt ggttattact gtctttactt ctaaagaatgt tagcttgcac tgaggatgaa 3360
 81 aagtgtgtac atatataata tacccttaca ttatgtatgc gggattttt taaattatata 3420
 82 taaaatgtct ccctagaatgc acaataggaa ggctaaatataa taataacctg tttctgggt 3480
 83 gttgtgggg catgagcttgc ttgtatcact gcttcataa actcaaccag ctgccttttt 3540
 84 aaagggagct ctatccctt ttgtgtattt cacttttattt attttatttac aaacttcaag 3600
 85 attatttaatgc tgaatattt tccatccatc tggggaaaat gccacagtgt tccatcg 3660
 86 gaacatccctt gcttgcgttgc aggctgtggg caagttcgttgc accacaggaa gtaaattggc 3720
 87 ctctttgata cacttttgc tgcctccca gaaaaaaagg aatttgcatttcc aaggatataca 3780
 88 tacatattca tcgtattttgc ttgtgttcc ttatgtatgc gggatttttt tttctgggt 3840
 89 tatactctgt gttctgttca tgcctctgttgc ttgttgcattt cttggatgat gagataggaa 3900
 90 aggagcaggatgc atgagacttgc caatggtcac agggaaaatgttgcattttt gtgtatgggtt 3960
 91 tattttctgt taacactgttgc ttgtgttgc ttatgtatgc gggatttttt tttctgggt 4020
 92 ctggatgggg gacagaaaaaa ggcagatcaccatc atgagatgttgc gggatttttt tttctgggt 4080
 93 ggcttacatgttgc ttgttgcattt tttttttttt tttttttttt tttttttttt tttttttttt 4140
 94 aaaaaaggatgc gggaaaatgttgc tttttttttt tttttttttt tttttttttt tttttttttt 4200
 95 ctatggccatc gaaaatattt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 4260
 96 gcacacttccatc cccttagatgc ccccttgcattt tttttttttt tttttttttt tttttttttt tttttttttt 4320
 97 caaaatgttgc aaataaaatgc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 4380
 98 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 4440
 99 tataaaaaatgc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 4500
 100 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 4560
 101 aaataaaatgc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 4588
 103 <210> SEQ ID NO: 2
 104 <211> LENGTH: 790
 105 <212> TYPE: PRT
 106 <213> ORGANISM: Human

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/003,608

DATE: 12/17/2001
TIME: 20:15:46

Input Set : A:\PTO.AMC.TXT
Output Set: N:\CRF3\12172001\J003608.raw

108 <400> SEQUENCE: 2
109 Met Ala Glu Gln Val Leu Pro Gln Ala Leu Tyr Leu Ser Asn Met Arg
110 1 5 10 15
111 Lys Ala Val Lys Ile Arg Glu Arg Thr Pro Glu Asp Ile Phe Lys Pro
112 20 25 30
113 Thr Asn Gly Ile Ile His His Phe Lys Thr Met His Arg Tyr Thr Leu
114 35 40 45
115 Glu Met Phe Arg Thr Cys Gln Phe Cys Pro Gln Phe Arg Glu Ile Ile
116 50 55 60
117 His Lys Ala Leu Ile Asp Arg Asn Ile Gln Ala Thr Leu Glu Ser Gln
118 65 70 75 80
119 Lys Lys Leu Asn Trp Cys Arg Glu Val Arg Lys Leu Val Ala Leu Lys
120 85 90 95
121 Thr Asn Gly Asp Gly Asn Cys Leu Met His Ala Thr Ser Gln Tyr Met
122 100 105 110
123 Trp Gly Val Gln Asp Thr Asp Leu Val Leu Arg Lys Ala Leu Phe Ser
124 115 120 125
125 Thr Leu Lys Glu Thr Asp Thr Arg Asn Phe Lys Phe Arg Trp Gln Leu
126 130 135 140
127 Glu Ser Leu Lys Ser Gln Glu Phe Val Glu Thr Gly Leu Cys Tyr Asp
128 145 150 155 160
129 Thr Arg Asn Trp Asn Asp Glu Trp Asp Asn Leu Ile Lys Met Ala Ser
130 165 170 175
131 Thr Asp Thr Pro Met Ala Arg Ser Gly Leu Gln Tyr Asn Ser Leu Glu
132 180 185 190
133 Glu Ile His Ile Phe Val Leu Cys Asn Ile Leu Arg Arg Pro Ile Ile
134 195 200 205
135 Val Ile Ser Asp Lys Met Leu Arg Ser Leu Glu Ser Gly Ser Asn Phe
136 210 215 220
137 Ala Pro Leu Lys Val Gly Gly Ile Tyr Leu Pro Leu His Trp Pro Ala
138 225 230 235 240
139 Gln Glu Cys Tyr Arg Tyr Pro Ile Val Leu Gly Tyr Asp Ser His His
140 245 250 255
141 Phe Val Pro Leu Val Thr Leu Lys Asp Ser Gly Pro Glu Ile Arg Ala
142 260 265 270
143 Val Pro Leu Val Asn Arg Asp Arg Gly Arg Phe Glu Asp Leu Lys Val
144 275 280 285
145 His Phe Leu Thr Asp Pro Glu Asn Glu Met Lys Glu Lys Leu Leu Lys
146 290 295 300
147 Glu Tyr Leu Met Val Ile Glu Ile Pro Val Gln Gly Trp Asp His Gly
148 305 310 315 320
149 Thr Thr His Leu Ile Asn Ala Ala Lys Leu Asp Glu Ala Asn Leu Pro
150 325 330 335
151 Lys Glu Ile Asn Leu Val Asp Asp Tyr Phe Glu Leu Val Gln His Glu
152 340 345 350
153 Tyr Lys Lys Trp Gln Glu Asn Ser Glu Gln Gly Arg Arg Glu Gly His
154 355 360 365
155 Ala Gln Asn Pro Met Glu Pro Ser Val Pro Gln Leu Ser Leu Met Asp
156 370 375 380

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/003,608

DATE: 12/17/2001
TIME: 20:15:46

Input Set : A:\PTO.AMC.TXT
Output Set: N:\CRF3\12172001\J003608.raw

157 Val Lys Cys Glu Thr Pro Asn Cys Pro Phe Phe Met Ser Val Asn Thr
158 385 390 395 400
159 Gln Pro Leu Cys His Glu Cys Ser Glu Arg Arg Gln Lys Asn Gln Asn
160 405 410 415
161 Lys Leu Pro Lys Leu Asn Ser Lys Pro Gly Pro Glu Gly Leu Pro Gly
162 420 425 430
163 Met Ala Leu Gly Ala Ser Arg Gly Glu Ala Tyr Glu Pro Leu Ala Trp
164 435 440 445
165 Asn Pro Glu Glu Ser Thr Gly Gly Pro His Ser Ala Pro Pro Thr Ala
166 450 455 460
167 Pro Ser Pro Phe Leu Phe Ser Glu Thr Thr Ala Met Lys Cys Arg Ser
168 465 470 475 480
169 Pro Gly Cys Pro Phe Thr Leu Asn Val Gln His Asn Gly Phe Cys Glu
170 485 490 495
171 Arg Cys His Asn Ala Arg Gln Leu His Ala Ser His Ala Pro Asp His
172 500 505 510
173 Thr Arg His Leu Asp Pro Gly Lys Cys Gln Ala Cys Leu Gln Asp Val
174 515 520 525
175 Thr Arg Thr Phe Asn Gly Ile Cys Ser Thr Cys Phe Lys Arg Thr Thr
176 530 535 540
177 Ala Glu Ala Ser Ser Ser Leu Ser Thr Ser Leu Pro Pro Ser Cys His
178 545 550 555 560
179 Gln Arg Ser Lys Ser Asp Pro Ser Arg Leu Val Arg Ser Pro Ser Pro
180 565 570 575
181 His Ser Cys His Arg Ala Gly Asn Asp Ala Pro Ala Gly Cys Leu Ser
182 580 585 590
183 Gln Ala Ala Arg Thr Pro Gly Asp Arg Thr Gly Thr Ser Lys Cys Arg
184 595 600 605
185 Lys Ala Gly Cys Val Tyr Phe Gly Thr Pro Glu Asn Lys Gly Phe Cys
186 610 615 620
187 Thr Leu Cys Phe Ile Glu Tyr Arg Glu Asn Lys His Phe Ala Ala Ala
188 625 630 635 640
189 Ser Gly Lys Val Ser Pro Thr Ala Ser Arg Phe Gln Asn Thr Ile Pro
190 645 650 655
191 Cys Leu Gly Arg Glu Cys Gly Thr Leu Gly Ser Thr Met Phe Glu Gly
192 660 665 670
193 Tyr Cys Gln Lys Cys Phe Ile Glu Ala Gln Asn Gln Arg Phe His Glu
194 675 680 685
195 Ala Lys Arg Thr Glu Glu Gln Leu Arg Ser Ser Gln Arg Arg Asp Val
196 690 695 700
197 Pro Arg Thr Thr Gln Ser Thr Ser Arg Pro Lys Cys Ala Arg Ala Ser
198 705 710 715 720
199 Cys Lys Asn Ile Leu Ala Cys Arg Ser Glu Glu Leu Cys Met Glu Cys
200 725 730 735
201 Gln His Pro Asn Gln Arg Met Gly Pro Gly Ala His Arg Gly Glu Pro
202 740 745 750
203 Ala Pro Glu Asp Pro Pro Lys Gln Arg Cys Arg Ala Pro Ala Cys Asp
204 755 760 765
205 His Phe Gly Asn Ala Lys Cys Asn Gly Tyr Cys Asn Glu Cys Phe Gln

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/003,608

DATE: 12/17/2001
TIME: 20:15:46

Input Set : A:\PTO.AMC.TXT
Output Set: N:\CRF3\12172001\J003608.raw

206	770	775	780
207	Phe Lys Gln Met Tyr Gly		
208	785	790	
211	<210> SEQ ID NO: 3		
212	<211> LENGTH: 1224		
213	<212> TYPE: DNA		
214	<213> ORGANISM: Human		
216	<220> FEATURE:		
217	<221> NAME/KEY: misc_feature		
218	<222> LOCATION: 36, 91, 645, 655, 660, 671, 672		
219	<223> OTHER INFORMATION: n = A,T,C or G		
221	<221> NAME/KEY: allele		
222	<222> LOCATION: (0)...(0)		
224	<400> SEQUENCE: 3		
W--> 225	tcgggatcga tctggagctc oggaaatttc cctggncgg gactccgggc tttccagccc 60		
W--> 226	caaccatgca taaaagggt tcggcgttct nggagagccca cagagccgg gccacaggca 120		
227	gctccttgc agctcttctt ctcctctcac agccgcaga cccgcctgct gagccccat 180		
228	ggcccgcgct gctctctccg cggcccccag caatccccgg ctcctgcgag tggcgctgct 240		
229	gctcctgctc ctggtagccg ctggccggcg cgccggcaggaa gcccggctgg ccactgaact 300		
230	gctgtccag tgcttgcaga ccctgcagg aattcaccc aagaacatcc aaagtgtgaa 360		
231	gttgaagtcc cccggacccc actgcgcaca aaccgaagtc atagccacac tcaagaatgg 420		
232	gcagaaaagct tgctcaacc cccgcacccatc catgtttaag aaaaatcatcg aaaagatgct 480		
233	gaaaaatggc aaatccaact gaccagaagg aaggaggaag cttattgtg gctgttcctg 540		
234	aaggaggccc tgcccttaca ggaacagaag aggaaagaga gacacagctg cagaggccac 600		
W--> 235	ctgggattgc gcctaattgtg tttgagcatc acttaggaga aggcncggat taatnaattn 660		
W--> 236	attatttat nnattgggtt gttttagaag atttatgtt aatatttat gtgtaaaata 720		
237	aggttatgat tgaatctact tgacacactt cccattatat ttattgttta ttttaggtca 780		
238	aacccaagtt agtcaatcc tgattcatat ttaatttcaa gatagaaggt ttgcagatat 840		
239	tctctagtca tttgttaata tttcttcgtg atgacatatac acatgtcagc cactgtgata 900		
240	gaggctgagg aatccaagaa aatggccagt aagatcaatg tgacggcagg gaaatgtatg 960		
241	tgtgtctatt ttgttaactgt aaagatgaat gtcagttgtt atttatttcaa atgatttcac 1020		
242	agtgtgtgtt caacatttctt catgttgaag ctttaagaac taaaatgttca taaatatccc 1080		
243	ttggacattt tatgtcttcc ttgttaaggca tactgcctt cttttttttt attatgcagt 1140		
244	gtttccctct gtgttagagc agagaggttt cgatattttt tgatgtttc acaaagaaca 1200		
245	ggaaaataaaa atattnaaaa atat 1224		
247	<210> SEQ ID NO: 4		
248	<211> LENGTH: 107		
249	<212> TYPE: PRT		
250	<213> ORGANISM: Human		
252	<400> SEQUENCE: 4		
253	Met Ala Arg Ala Ala Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu 15		
254	1 5 10 15		
255	Arg Val Ala Leu Leu Leu Leu Leu Val Ala Ala Gly Arg Arg Ala 20 25 30		
256	20 25 30		
257	Ala Gly Ala Pro Leu Ala Thr Glu Leu Arg Cys Gln Cys Leu Gln Thr 35 40 45		
258	35 40 45		
259	Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Lys Val Lys Ser 50 55 60		
260	50 55 60		
261	Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn		

fwj Use of n and/or Xaa has been detected in the Sequence Listing.
Review the Sequence Listing to insure a corresponding explanation is presented in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/003,608

DATE: 12/17/2001
TIME: 20:15:47

Input Set : A:\PTO.AMC.TXT
Output Set: N:\CRF3\12172001\J003608.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:225 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:226 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:235 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:236 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:476 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11